The precision of identified variables with respect to multivariable set size in glycaemic data from a virtual type 1 diabetic patient

Erin J. Mansell, Paul D. Docherty and J. Geoffery Chase

Department of Mechanical Engineering, University of Canterbury, Christchurch, Private Bag 4800, New Zealand
(Tel: +64 3 3642987; e-mail:erin.mansell@pg.canterbury.ac.nz).

Abstract: Prior research had been carried out to identify a large number of glycaemic variables in sparse, noisy data from a virtual diabetic patient. This paper investigates the precision of variables as an identification scheme introduces progressively more parameters into the variable set and as the quantity of data increases. Virtual data was simulated with a diabetic glycaemic meal model that contained six variable parameters. Data was sampled 6 times daily with noise. Increasing variable sets were identified for data subsets of increasing size. Norm-error of equivalent variable groups was compared before and after new parameter introductions. A Monte Carlo analysis was carried out to evaluate a population of results.

Identifying new variables improved parameter estimates in all equivalent variable groups by 34 days in the mean population case. However, variability from data noise resulted in some cases never yielding six-parameter identification that improved upon results that relied on *a-priori* information. When parameters were introduced as variables too soon for the given data quality/quantity, reduced practical identifiability caused interference between these and other variables, diminishing their precision. However, when introduced too late the precision in the variable set was hindered by effects not fully described by the *a-priori* guesses. Introducing the 3rd and 4th variables early in the data produced significant benefit in most cases. In contrast, the 5th and 6th parameters could not be introduced as early, improved precision by a lesser degree on average and in many cases never improved precision. The influence of noise on practical identifiability highlighted the need for similar analyses *in-vivo* so as to strategise parameter identification to gain the most information at the highest precision.

Keywords: parameter identification, identifiability, physiological models, Monte Carlo simulation, type 1 diabetes, stress hyperglycaemia, exercise

1. INTRODUCTION

Individuals with type 1 diabetes mellitus take regular doses of insulin to overcome an almost complete lack of pancreatic insulin production. The goal of insulin therapy is to minimise postprandial hyperglycaemia, while maintaining low risk of hypoglycaemia (Rubin and Peyrot, 2001). Both states are associated with reduced quality of life and a number of serious health complications (Rubin and Peyrot, 1999, De Boer *et al.*, 2008, Retnakaran and Zinman, 2008, Shankar *et al.*, 2007).

The optimal insulin dose for a given meal depends largely on the carbohydrate content and the insulin sensitivity (*SI*) of the patient, but many smaller and unquantified influences are also in effect, which can add significant complication to self-managed blood glucose (BG) control of the most diligent individuals. Stress, fatigue, and circadian metabolic rhythms are factors that modify the effective *SI* (Surwit *et al.*, 1992, Räikkönen *et al.*, 1996, Lloyd *et al.*, 1999). Furthermore, exercise has a significant effect on the rate of glucose disappearance (Sonnenberg *et al.*, 1990, Roy and Parker, 2007, Yardley *et al.*, 2013). The ability to quantify these unmeasured 'secondary effects' for specific individuals has the potential to improve self-regulated normo-glycaemia.

In prior research we identified seven glycaemic parameters in sparse BG data from a virtual patient (Mansell, *et al.*, unpublished). The results of the Monte Carlo analyses proved the parameters were all observable, and that measurement noise and un-modelled effects could be overcome as data size increased, with coefficient of variation (CV) across the population reducing in proportion to $1/\sqrt{n}$ and only small (<1%) biases resulting after one year (Mansell *et al.*, unpublished). To observe long-term drift in *SI*, a 90 day data window was progressed along a 2 year period to identify parameters in a moving average. Information for stress, fatigue and exercise parameters were still able to be captured with biases remaining less than 1% and CV being approximately equivalent to drift-free outcomes.

Practical non-identifiability occurs when experimental data is of insufficient quality or quantity compared to the size of a model (Raue *et al.*, 2009, Docherty *et al.*, 2011). This paper investigates identifiability in the six most variable parameters

from the original seven-parameter cohort. Specifically, we aim to determine the data quantity required to support all six parameters without encountering such parameter interference.

2. METHODS

All analysis used MATLAB R2014a.

2.1 Simulating virtual patient model

Glycaemic dynamics for the virtual patient were modelled using a variation of DISST (Lotz *et al.*, 2010) and nutrition models (Wong *et al.*, 2008a, Wong *et al.*, 2008b). These were further modified to include secondary effect parameters for stress, fatigue, exercise and circadian rhythms of *SI*. The first order, multi-compartmental model consisted of many constant parameters (definitions in Table 1) and time-dependent inputs (Table 2), some of which were randomised in occurrence and/or magnitude for each repetition of the patient for the Monte Carlo analysis.

Subcutaneous insulin (U_S) absorbed from regular insulin boluses (U_X) was modelled with the analytical solution to:

$$\dot{U}_S(t) = k_X(U_X(t) - U_S(t)) \quad (1)$$

Interstitial insulin concentration (Q), dependant on U_S and plasma insulin (I), was simulated with a true analytical solution to Equations 2 and 3 found with MATLAB's symbolic differential equation solver (dsolve) assuming the initial state is at equilibrium:

$$\dot{Q}(t) = -(n_I + n_C)Q(t) + n_I I(t) \quad (2)$$

$$\dot{I}(t) = -(n_T + n_I)I(t) + n_I Q(t) + k_X U_S(t)/V_P \quad (3)$$

Glucose absorbed into the gut (P_S) was calculated from regular ingested meals of varying glucose content (P_X) and randomly timed snacks (P_C):

$$\dot{P}_S(t) = \big(P_X(t) + P_C(t)\big)/V_G - k_1 P_S(t) \quad (4)$$

Insulin sensitivity (SI) was modelled as a the sum of overlapping triangular basis functions ($g_{1\text{-}3}$) multiplied by their corresponding morning, midday and afternoon peak insulin sensitivities ($SI_{1\text{-}3}$), along with stress (σ) and fatigue (φ) modifiers:

$$SI(t) = \big(SI_1 g_1(t) + SI_2 g_2(t) + SI_3 g_3(t)\big) \times \big(1 - \sigma(t) - \varphi(t)\big) \quad (5)$$

The basis functions and SI profile are pictured in Figure A1, Appendix A, for clarification.

BG (G) was modelled using Q, P_S and SI, along with an effect from exercise (ε):

$$\dot{G}(t) = k_2 P_s(t) - SI(t)(G(t)\cdot Q(t) - G_0 Q_0) - p_G\big(G(t) - G_0 + \varepsilon(t)\big) \quad (6)$$

Each of the stress, fatigue, and exercise functions was calculated as the product of a maximum value and a time-dependent function with effect intensity ranging from 50-100% for each appearance of the effect and 0% otherwise.

$$\varepsilon(t) = \varepsilon_{max}\cdot f_\varepsilon(t), \quad \sigma(t) = \sigma_{max}\cdot f_\sigma(t), \quad \varphi(t) = \varphi_{max}\cdot f_\varphi(t) \quad (7.a\text{-}c)$$

All time-dependent variables were simulated to one minute resolution. Definite integrals were calculated with trapezoidal numerical integration.

2.2 Simulation and data sampling

Diary-like data was created for the virtual patient where daily finger-prick measurements were documented and the carbohydrate content of meals estimated, neglecting 'unrecorded' snacks. Insulin doses were recorded, as well as instances and intensity of exercises, stress and fatigue.

Table 1. Parameter constants used to simulate the virtual patient glycaemic profiles in Equations 1-6

Parameter	Description	Value	Unit
n_I	Plasma to interstitium transport rate	0.02	min^{-1}
n_T	Plasma insulin clearance rate	0.1	min^{-1}
n_C	Cell metabolism of insulin	0.02	min^{-1}
v_P	Volume of distribution of plasma insulin	4.3	L
p_G	Glucose dependant balance	0.004	min^{-1}
V_G	Glucose distribution volume	12.4	L
k_1	Rate of glucose transfer from stomach to gut	0.05	min^{-1}
k_2	Rate of glucose absorption from gut	0.008	min^{-1}
k_X	Rate of insulin dispersed from injection site	0.01	min^{-1}
G_0	Basal glucose level	4.5	$mmol.L^{-1}$
Q_0	Basal interstitial insulin level	4.23	$mU.L^{-1}$
ε_{max}	Exercise coefficient	6.5	$mmol.L^{-1}$
σ_{max}	Stress coefficient	0.3	
φ_{max}	Fatigue coefficient	0.1	
SI_1	Morning (8.30am) peak	$0.8\text{x}10^{-3}$	$L.mU^{-1}.min^{-1}$
SI_2	Midday (12pm) peak	$1.0\text{x}10^{-3}$	$L.mU^{-1}.min^{-1}$
SI_3	Afternoon (3.30pm) peak	$0.6\text{x}10^{-3}$	$L.mU^{-1}.min^{-1}$

Table 2. Time-dependent vector inputs used in patient simulation with one minute resolution

Input	Description	Value	Unit
P_X	meals	[*400*,*500*] at 0800, 1200 and 1900 hrs daily, *0* otherwise	mmol
P_C	snacks	*160* at 52 random *t* per year, *0* otherwise	mmol
U_X	insulin doses	*1000* with meals, *4* otherwise	mU.min^{-1}
f_ε	exercise intensity	∈[*0.5, 0.6, ..., 1.0*] at 0830 to 1030 hrs, 3 days/week, *0* otherwise	
f_σ	stress intensity	∈[*0.5, 0.6, ..., 1.0*] 3 days per 4 weeks, *0* otherwise	
f_φ	fatigue intensity	∈[*0.5, 0.6, ..., 1.0*] 5 days per 4 weeks, *0* otherwise	
g_1	morning *SI* basis	triangular function from *0* at 1530 hrs to *1* at 0830 to *0* at 1200 hrs	
g_2	midday *SI* basis	triangular function from *0* at 0830 hrs to *1* at 1200 to *0* at 1530 hrs	
g_3	afternoon *SI* basis	triangular function from *0* at 1200 hrs to *1* at 1530 to *0* at 0830 hrs	

In practice, to simulate the data required for parameter identification, 6 randomly-timed samples were taken from *G(t)* between 6am and 12 midnight daily with 10% normally distributed multiplicative white noise. A different $P_S(t)$ was calculated neglecting P_C and with 10% uniformly distributed noise applied to P_X. All remaining parameters not included in any identification set were treated as in Tables 1-2.

2.3 Parameter identification

The Gauss-Newton (GN) method of gradient descent was used to identify the least squares solution of variable set **x** by minimising the residual error between the sampled and forward-simulated BG over several iterations:

$$\boldsymbol{\psi}_i = \mathbf{G}_i(t_S, \mathbf{x}_i) - \mathbf{G_S} \qquad (7)$$

where $\mathbf{G}_S$ denotes the sampled data and $\mathbf{G}_i(t_S, \mathbf{x}_i)$ denotes the modelled glucose concentration at the sample times (t_S) and the present iteration (i) using the current variable set. *S* denotes the samples 1…*n*, where *n* is the number of samples.

Parameters of interest were subsets of those in the set:

$$\mathbf{x} = [SI_1, SI_2, SI_3, \varepsilon_{max}, \sigma_{max}, \varphi_{max}]^{\mathrm{T}} \qquad (8)$$

where between 2 and all 6 of these parameters were identified as variables when required. Variables were initially set to relevant subsets of $\mathbf{x}_0 = [10^{-3}, 10^{-3}, 10^{-3}, 1, 0.1, 0.1]^{\mathrm{T}}$ and GN iterations were continued until the tolerance criteria $||(\mathbf{x}_{i-1} - \mathbf{x}_i)/\mathbf{x}_0||_2 < 10^{-4}$ yielded approximately 4 significant figures of convergence precision on parameter estimates.

2.4 Structural identifiability and stability checks

When instances of exercise, stress or fatigue were not present in an identified period of time, ε_{max}, σ_{max} and φ_{max} were set to zero and excluded from identification. When both stress and fatigue effects occurred concurrently yielding a structurally non-identifiable system, both parameters were excluded, as their effects can only be quantified when there are distinguishable instances (Docherty *et al.*, 2011, Bellu *et al.*, 2007).

For occasional instances in small sets of data where noise and un-modelled effects rendered GN identification unstable (when singular matrix occurred or $i > 30$), latter variables were removed from the process every 20 iterations until successful identification occurred.

2.5 Error analysis

Proof of concept analysis was carried out as follows:

1. 40 days of data was simulated
2. This data was broken into subsets of 0-1,0-2,…,0-40 days
3. $SI_{1\text{-}2}$ were identified for all data sets while other variables were set as '*a-priori*' parameters using incorrect values: $SI_3 = 0.4\text{x}10^{-3}$, $\varepsilon_{\max} = \sigma_{\max} = \varphi_{\max} = 0$
4. SI_3 was introduced to the variable set and identification was again carried out over all data sets
5. Error in SI_1 and SI_2 was compared for the 2-variable and 3-variable cases. Norm-error was calculated as $e = \left\| \frac{\mathbf{x} - \mathbf{x}_{true}}{\mathbf{x}_{true}} \right\|_2$ where $\mathbf{x}_{\text{true}}$ is the original model input
6. Beneficial data size at which to introduce SI_3 was defined as when norm-error of $SI_{1\text{-}2}$ was less for the 3-variable case than 2-variable case

2.6 Monte Carlo analysis

Population-wide analysis employed the following process:

1. 89 days of data was simulated for 1000 patients
2. These data sets were broken into subsets from zero to [2, 3, 5, 8, 13, 21, 34, 55, 89] days
3. $SI_{1\text{-}2}$ were identified for all data subsets for all repeats
4. SI_3, $\varepsilon_{\max}$, $\sigma_{\max}$, and $\varphi_{\max}$ were all added to the variable set one at a time, with new identification occurring at each addition
5. Norm-error was calculated for variable subsets $x_{1\text{-}2}$ for all results, $x_{1\text{-}3}$ when 3 or more variables were identified, and so on up to norm-error of $x_{1\text{-}5}$ for 5 and 6 variable cases
6. Days when an increase in the number of variables yielded improved precision in each previous variable subset were located for the population mean
7. For the 89 day subset only, mean reduction in norm-error for equivalent variable subsets was calculated for each variable introduction, defined as $1 - \bar{e}_{new}/\bar{e}_{old}$. The fraction of non-improved cases was also calculated

3. RESULTS

Figure 1 shows the *SI* parameter estimates when the identification set is 2 variables vs 3 variables. Note that the 3-variable set yielded larger error in SI_2 for days 1-10 compared to the 2-variable set. However, after approximately 23 days, the combined norm-error for $SI_{1\text{-}2}$ is reduced with the introduction of SI_3 into the variable set, Figure 2. For the MC population mean, the point at which a 3-variable set favoured the precision of $SI_{1\text{-}2}$ was at four days, Figure 3. Further introduction of ε_{max}, σ_{max}, and φ_{max} variables reduced $SI_{1\text{-}2}$ norm-error at three, nine and 34 days, Figure 4

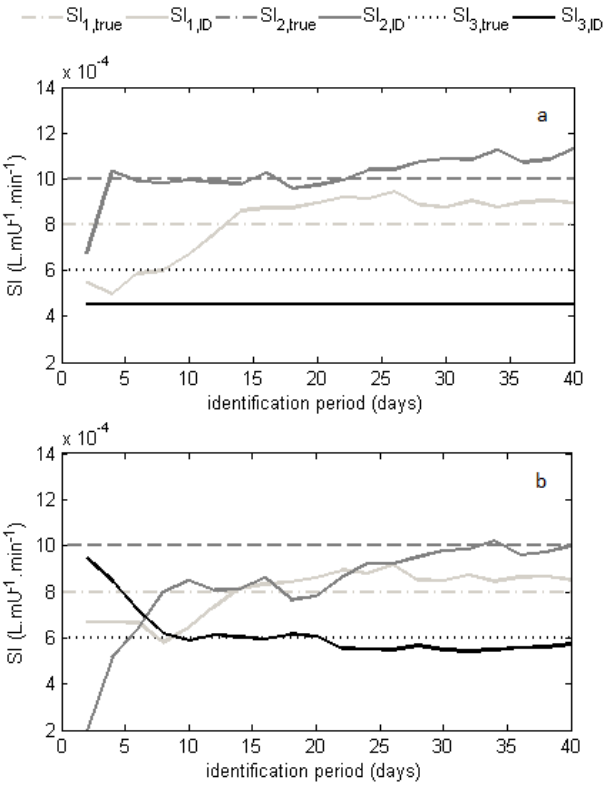

Figure 1. Identification for a data set with 2 variables plus one *a-priori* parameter (a) and all 3 as variables (b). Note: the legend is located at the top of the figure.

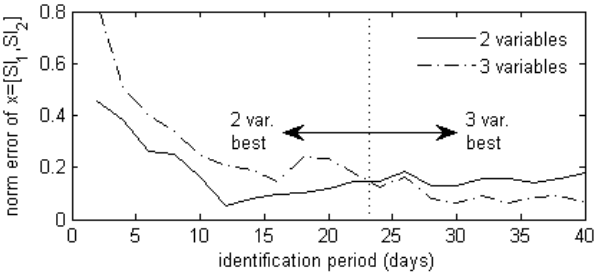

Figure 2. Norm-error of $SI_{1\text{-}2}$ for 2 and 3 variable identification sets. Prior to 23 days the least error is achieved by treating SI_3 as *a-priori*, afterward it is better identified as a variable.

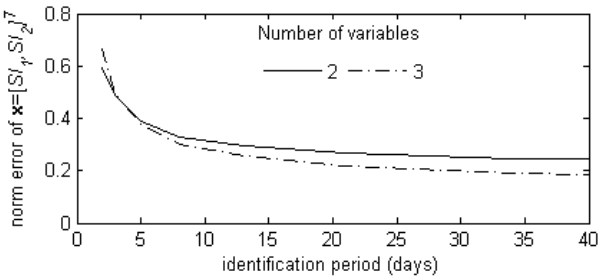

Figure 3. Norm-error of $SI_{1\text{-}2}$ as 2 and 3 variables are identified over a population, reduction in error for 3 variables at 4 days.

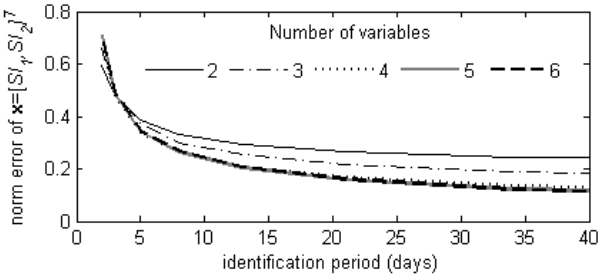

Figure 4. Norm error of $SI_{1\text{-}2}$ as 2-6 variables are identified over a population. Reduced error for variable introductions 3-6 occurs at 4, 3, 9 and 34 days.

Further consideration was taken into to the norm-error of variable sets greater than just $SI_{1\text{-}2}$.

Table 1 shows that for the population mean, crossover days for beneficially introducing certain parameters into the variable set were comparable across all norm-error sets (e.g. ε_{max} improves norm-error of both $SI_{1\text{-}2}$ and $SI_{1\text{-}3}$ after 3 days). Additionally, the crossover days generally increased for subsequent variable additions (e.g. the 5th parameter, σ_{max}, can be introduced as a variable on average 3 days after the 4th parameter, ε_{max}).

Table 2 gives statistics for the degree of norm-error improvement after 89 days for each increase in variable number. Identifying SI_3 reduced norm-error in $SI_{1\text{-}2}$ by 29% on average, only failing to improve the error in 1% of cases. Introducing the exercise parameter, ε_{max}, produced greater mean error reductions (28 and 41% for $SI_{1\text{-}2}$ and $SI_{1\text{-}3}$), but also had a greater no-benefit rate of 6-18%. Introducing stress, σ_{max}, and fatigue, φ_{max}, as variables had less benefit in error reduction, especially for φ_{max}, where mean reductions ranged from an increase in error of 6% to a reduction of 6% and no-benefit rates were 28-40%.

Table 1. Mean day for an introduced parameter to improve precision in other variable sets.

parameter introduced	variable subset evaluated				
	$x_{1\text{-}2}$	$x_{1\text{-}3}$	$x_{1\text{-}4}$	$x_{1\text{-}5}$	average
$x_3 = SI_3$	4				4
$x_4 = \varepsilon_{max}$	3	3			3
$x_5 = \sigma_{max}$	9	5	5		6
$x_6 = \varphi_{max}$	34	27	27	26	29

Table 2. Population variability statistics for 89 days of data with mean reduction in norm-error for parameter sets as subsequent parameters were introduced and the no-benefit rate where precision was not improved.

parameter introduced	statistic	variable subset evaluated			
		$x_{1\text{-}2}$	$x_{1\text{-}3}$	$x_{1\text{-}4}$	$x_{1\text{-}5}$
$x_3 = SI_3$	reduction	29%			
	no-benefit	1.3%			
$x_4 = \varepsilon_{max}$	reduction	28%	41%		
	no-benefit	18%	6.2%		
$x_5 = \sigma_{max}$	reduction	9.0%	23%	16%	
	no-benefit	20%	8.6%	16%	
$x_6 = \varphi_{max}$	reduction	-6.2%	6.2%	3.7%	2.6%
	no-benefit	40%	28%	34%	36%

4. DISCUSSION

Introducing a new parameter into the identified set of variables can reduce the error of the original variables due to the ability of the introduced parameter to assume a value that has less associated error than an *a-priori* guess. A comparison of Figure 1a and 1b shows that by the end of 40 days, the error in $SI_{1\text{-}3}$ was less when SI_3 was identified as a variable, rather than taken as an incorrect *a-priori* parameter. However, too little data results negates this beneficial effect. Figure 1b shows that identifying SI_3 introduced large error in SI_2 for small n. This outcome demonstrates that, at first, the data was not sufficient to support all the variables with any degree of accuracy, and reduced practical identifiability resulted in variable interference (Raue *et al.*, 2009, Docherty *et al.*, 2011).

$SI_{1\text{-}3}$ are equivalent parameters that peak at different times of the day. Hence the times at which data points were taken would influence accuracy in $SI_{1\text{-}3}$ estimates. For example, if data points have largely been sampled in the morning, then the effect of the afternoon *SI* is not easily distinguishable, and the resulting error may propagate particularly strongly into the midday *SI*, as was the likely case in Figure 1b. As more data accumulated, with more random additions inevitably occurring in the afternoon, the afternoon *SI* began to achieve accuracy and all variables were benefited. This illustrates the concept of practical identifiability well.

While mean population results appear to indicate cleanly when parameters should be introduced as variables (Figure 4), the trends fail to capture variability effects and thus represent an ideal case rather than average. In particular the mean averages out error, the y-dimension property in Figure 4, but not the crossover day when the variable set can be increased, the x-dimension property. In fact, for individual sets of data there were frequently multiple crossovers points or none at all in the first 89 days. This reality cannot be captured by the mean, thus the statistical data of Table 2 was calculated.

Based on results for variability effects, SI_3 appears beneficial to introduce by day 89 since it reduced norm-error of $SI_{1\text{-}2}$ in 99% of cases. Since the appearance of SI_3 was daily, its accuracy weighed heavily on the outcomes of other variables. Comparatively, stress occurred much less frequently and had lower gains for greater risk (no-benefit rate 9-20%). Exercise, $\varepsilon_{max,}$ also appears beneficial to introduce with large benefits (28-41%) though a moderate no-benefit rate (6-18%). Since the model includes frequent exercise at the same time of day, not identifying ε_{max} is likely to skew at least one *SI* peak value, thus the risk could be deemed acceptable compared to the benefits. Like stress, fatigue was relatively infrequent. Thus, many more days of data were required to achieve accuracy and minimise interference with other variables. This can be seen in the higher no-benefit rates (28-40%) and lower gains (-6% to 6% error reduction). Perhaps both σ_{max} and φ_{max} should not be identified due to their tendency for error propagation.

Of course, while introducing some parameters is likely to increase error in other parameters, this detriment must be weighed against the benefit identifying the new parameter itself. A small error increase could be an acceptable price for the advantageous information. However, a best case for all parameters could be achieved by identifying a base parameter set of regular and highly identifiable parameters, then fixing some or all of these parameters while performing a second identification round to ascertain the less frequent parameters.

An *in-silico* analysis was the best platform for investigating the research presented in this paper since true parameters values are non-existent in real data. Therapeutic glycaemic modelling can be difficult in practice because of the presence of noise, un-modelled effects and sometimes practical non-identifiability. We have shown that much of these effects can be accounted for in a stable manner through the timely addition of new parameters into the identified set of variables. It remains unknown how real data would respond to equivalent parameter introductions. However, the concepts explored act as a starting point to developing and testing the efficacy of other analysis methods independent of error and perhaps instead evaluating properties such as variability.

There are many factors that can affect the glycaemic dynamics of people with diabetes to some tangible degree, not limited to those modelled in this report. The ability to identify a large number of such parameters in one set of data would be valuable. However, if certain parameters are introduced too soon in the accumulation of data, they can seriously reduce the precision of the other parameter estimates. If introduced too late, then un-modelled behaviour, or 'grey' noise, is the limiting factor on the precision of identified variables. Hence the importance of this type of identifiability analysis.

5. CONCLUSIONS

Identification of increasing numbers of parameters generally improves error in the parameter group by capturing otherwise un-modelled effects. However insufficient data can reduce practical identifiability, increasing parameter interference and error. The point at which data does become sufficient was diagnosed through analysis of error in equivalent variable sets before and after parameter introduction.

Specific to this model and analysis, the two parameters SI_3 and ε_{max} appear reasonably beneficial to introduce after 89 days, while σ_{max} and φ_{max} may be better excluded until the other parameters can be well established and fixed.

Error-based analysis was ideal for exploring the effect of practical identifiability on the model, but will ultimately be ineffective for *in vivo* data, requiring improved methods. Identifying large numbers of parameters with known confidence would be useful in diabetes. Greater knowledge of secondary glycaemic factors, could achieve improved glycaemic control with greater lifestyle flexibility.

REFERENCES

Bellu, G., Saccomani, M. P., Audoly, S. & D'angiò, L. (2007). DAISY: A new software tool to test global identifiability of biological and physiological systems. *Computer Methods and Programs in Biomedicine,* 88**,** 52-61.

De Boer, I. H., Kestenbaum, B., Rue, T. C., Steffes, M. W., Cleary, P. A., Molitch, M. E., Lachin, J. M., Weiss, N. S. & Brunzell, J. D. (2008). Insulin therapy, hyperglycemia, and hypertension in type 1 diabetes mellitus. *Archives of Internal Medicine,* 168**,** 1867-1873.

Docherty, P. D., Chase, J. G., Lotz, T. F. & Desaive, T. (2011). A graphical method for practical and informative identifiability analyses of physiological models: A case study of insulin kinetics and sensitivity. *BioMedical Engineering Online,* 10.

Lloyd, C. E., Dyer, P. H., Lancashire, R. J., Harris, T., Daniels, J. E. & Barnett, A. H. (1999). Association between stress and glycemic control in adults with type 1 (insulin-dependent) diabetes. *Diabetes Care,* 22**,** 1278-1283.

Lotz, T. F., Chase, J. G., Mcauley, K. A., Shaw, G. M., Docherty, P. D., Berkeley, J. E., Williams, S. M., Hann, C. E. & Mann, J. I. (2010). Design and clinical pilot testing of the model-based Dynamic Insulin Sensitivity and Secretion Test (DISST). *Journal of Diabetes Science and Technology,* 4**,** 1408-1423.

Mansell, E. J., Docherty, P. D., Fisk, L. M. & Chase, J. G. (unpublished). The observability of secondary effects on glycaemic dynamics in accumulating data from a virtual type 1 diabetic patient. *Journal of Mathematical Biosciences*.

Räikkönen, K., Keltikangas-Järvinen, L., Adlercreutz, H. & Hautanen, A. (1996). Psychosocial stress and the insulin resistance syndrome. *Metabolism: Clinical and Experimental,* 45**,** 1533-1538.

Raue, A., Kreutz, C., Maiwald, T., Bachmann, J., Schilling, M., Klingmüller, U. & Timmer, J. (2009). Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics,* 25**,** 1923-1929.

Retnakaran, R. & Zinman, B. (2008). Type 1 diabetes, hyperglycaemia, and the heart. *The Lancet,* 371**,** 1790-1799.

Roy, A. & Parker, R. S. (2007). Dynamic Modeling of Exercise Effects on Plasma Glucose and Insulin Levels. *Journal of Diabetes Science and Technology,* 1**,** 338-347.

Rubin, R. R. & Peyrot, M. (1999). Quality of life and diabetes. *Diabetes/Metabolism Research and Reviews,* 15**,** 205-218.

Rubin, R. R. & Peyrot, M. (2001). Psychological issues and treatments for people with diabetes. *Journal of Clinical Psychology,* 57**,** 457-478.

Shankar, A., Klein, R., Klein, B. E. K. & Moss, S. E. (2007). Association between glycosylated hemoglobin level and cardiovascular and all-cause mortality in type 1 diabetes. *American Journal of Epidemiology,* 166**,** 393-402.

Sonnenberg, G. E., Kemmer, F. W. & Berger, M. (1990). Exercise in type 1 (insulin-dependent) diabetic patients treated with continuous subcutaneous insulin infusion. Prevention of exercise induced hypoglycaemia. *Diabetologia,* 33**,** 696-703.

Surwit, R. S., Schneider, M. S. & Feinglos, M. N. (1992). Stress and diabetes mellitus. *Diabetes Care,* 15**,** 1413-1422.

Wong, J., Chase, J. G., Hann, C. E., Shaw, G. M., Lotz, T. F., Lin, J. & Le Compte, A. J. (2008a). A Subcutaneous Insulin Pharmacokinetic Model for Computer Simulation in a Diabetes Decision Support Role: Model Structure and Parameter Identification. *Journal of Diabetes Science and Technology,* 2**,** 658-671.

Wong, J., Chase, J. G., Hann, C. E., Shaw, G. M., Lotz, T. F., Lin, J. & Le Compte, A. J. (2008b). A Subcutaneous Insulin Pharmacokinetic Model for Computer Simulation in a Diabetes Decision Support Role: Validation and Simulation. *Journal of Diabetes Science and Technology,* 2**,** 672-680.

Yardley, J. E., Sigal, R. J., Perkins, B. A., Riddell, M. C. & Kenny, G. P. (2013). Resistance exercise in type 1 diabetes. *Canadian Journal of Diabetes,* 37**,** 420-426.

Appendix A. VISUAL CLARIFICATION OF INSULIN SENSITIVITY PROFILE

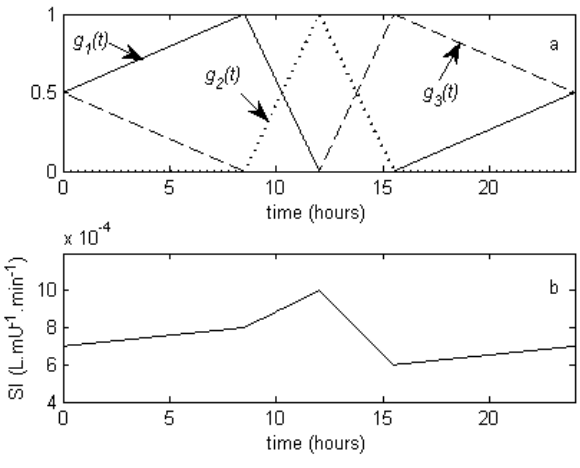

Figure A1. The triangular basis functions $g_{1\text{-}3}$ (a) used to achieve the *SI* profile (b) with Equation 6.